Request for Reconsideration after Final Action

The table below presents the data as entered.

Input Field	Entered
SERIAL NUMBER	85561168
LAW OFFICE ASSIGNED	LAW OFFICE 106
MARK SECTION (no change)	
ARGUMENT(S)	

Applicant appreciates the acknowledgment by the Examining Attorney that the amendment to the services listed for International Class 042 and substitute specimen submitted for International Class 042 have satisfied requirement of the earlier office action.

The office action maintains that there is a likelihood of confusion between Applicant's mark and the cited mark because the term WEB INTELLIGENCE is the dominant feature of the two marks and the goods sold under the marks are related. Applicant respectfully requests reconsideration.

The office action indicates that there is potential confusion between the cited mark and Applicant's mark because the goods sold under the two marks are related. The Examining Attorney indicated that the goods associated with the cited mark overlap those of Applicant's goods because the Registrant's goods are not limited to the "surface web." However, the Registrant's goods are limited to "information stored in relational databases," which is the same as the "surface web."

The Examining Attorney also maintains that Applicant's mark of DEEP WEB INTELLIGENCE is merely Registrant's WEB INTELLIGENCE mark with the simple addition of a descriptive term. However, as previously presented by Applicant and further clarified below, the customers of both the Registrant and Applicant understand that the term DEEP WEB has a completely different connotation than WEB with a simple modifier.

A. Applicant's Goods and Services Are Distinct From the Goods of the Cited Mark.

While Applicant's goods and Registrant's goods may be complementary or related, the Office Action misses a significant point. Applicant's goods and services are expressly limited to the "Deep Web." Registrant's goods are expressly limited to "relational databases." These areas of the internet are exclusive to one another. Applicant's goods and services are used for searching data not found in relational databases.

Attached is a whitepaper written by the CEO of Applicant published in March 2013 that explains the significant difference between searching relational databases versus searching the Deep Web. Contrary to the assertion in the office action that the Registrant's mark is not limited to the "surface web," the

goods are expressly limited "relational databases," i.e., structured content or, as previously noted, "the surface web."

The goods and services of the parties need not be identical or directly competitive to find a likelihood of confusion. However, they need be related in some manner, or the conditions surrounding their marketing be such, that they could be encountered by the same purchasers under circumstances that could give rise to the mistaken belief that the goods and services come from a common source. *In re Martin's Famous Pastry Shoppe, Inc.*, 748 F.2d 1565, 223 U.S.P.Q. 1289 (Fed. Cir. 1984); *In re Corning Glass Works*, 229 U.S.P.Q. 65 (T.T.A.B. 1985); *In re Rexel Inc.*, 223 U.S.P.Q. 830 (T.T.A.B. 1984); *Guardian Products Co., Inc. v. Scott Paper Co.*, 200 U.S.P.Q. 738 (T.T.A.B. 1978); *In re International Telephone & Telegraph Corp.*, 197 U.S.P.Q. 910 (T.T.A.B. 1978). As outlined in the attached whitepaper, Applicant's goods and services are utilized by military, intelligence, and law enforcement agencies for mining information beyond that available via search engines used with Registrant's goods. Therefore, it is submitted that confusion is unlikely.

B. Cited Marks are Not Distinctive.

Another factor that supports the fact that there is no likelihood of confusion is the fact that there are several marks using the term WEB that have been deemed allowable over each other. Attached is the first page of the results for current registrations using the term WEB in International Class 009 showing there are 724 such registrations. As stated by J. Thomas McCarthy, the noted trademark authority in his well-known treatise Trademarks and Unfair Competition, Fourth Edition, Volume 2, § 11.85:

> A mark that is hemmed in on all sides by similar marks on similar goods cannot be very "distinctive." It is merely one of a crowd of marks. In such a crowd, customers will not be likely to be confused between any two of the crowd. In such a crowd, customers will not likely be confused between any two of the crowd and may have learned to carefully pick out one from the other.

Not only are there numerous registrations in International Class 009 for WEB, but there are examples of what the Examining Attorney contends is a "modifier" used with another mark. Attached are summaries of U.S. Registration No. 3642023 for THE SOCIAL WEB BROWSER and U.S. Registration No. 4303507 for WEB BROWSER, both in International Class 009.

The basic factor in the analysis is whether the relevant purchasing public is likely to be confused by the use of the Applicant's mark. Whereas, herein, there are a number of similar marks for related goods, that the public is not likely to associate the Applicant's product bearing the mark WEB INTELLIGENCE with the Registrant's mark or the other WEB above.

C. Applicant's and Registrant's Respective Customers are Sophisticated and Discerning .

In further support of registration, Applicant notes that the sophistication of the buyer is an important factor in the likelihood of confusion analysis. In *Mead Data Central. Inc. v. Toyota Motor Sales USA. Inc.,* 9 U.S.P.Q. 2d 1442, 1449 (S.D. N.Y.), the Court elaborated on the importance of this factor to the likelihood of confusion analysis, saying that "typically, sophistication of the buyer is a factor that will weigh against finding a likelihood of confusion." As noted above, Applicant's customers are sophisticated, informed purchasers of the military, intelligence, and law enforcement community seeking information that is not found on the surface web. Registrant's customers seek business data in relational databases (see attached brochure of Registrant's product). It is likely that consumers of companies that use relational database searching would not be involved in unstructured data searching

or, if so, would understand the differences in such services.

Given the parties' respective clientele, it is clear that confusion is not likely because Applicant's customers are sophisticated enough to distinguish between Applicant's mark and Registrant's mark. Further, Applicant's mark and the cited mark are not likely to be encountered by the same customers because of the Applicant's specialty field of trade. Considering these facts, Applicant submits that confusion is not likely.

The basic factor in the analysis is whether the relevant purchasing public is likely to be confused by the use of Applicant's mark. Due to the sophistication of Applicant's and Registrant's customers and the exclusivity of the goods for which the marks apply, there is no likelihood of confusion.

Applicant accordingly submits that the objections should be withdrawn and its mark approved for publication. The Examiner is invited to telephone the undersigned if the Examiner believes it would be useful to advance prosecution.

EVIDENCE SECTION

EVIDENCE FILE NAME(S)	
ORIGINAL PDF FILE	evi_207250239110-120501607_._Deep_Web_Whitepaper.pdf
CONVERTED PDF FILE(S) (11 pages)	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0002.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0003.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0004.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0005.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0006.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0007.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0008.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0009.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0010.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0011.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0012.JPG
ORIGINAL PDF FILE	evi_1-207250239110-120501607_._List_of_WEB_Registrations.pdf
CONVERTED PDF FILE(S) (5 pages)	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0013.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0014.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0015.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0016.JPG

	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0017.JPG
ORIGINAL PDF FILE	evi_2-207250239110-120501607_._Web_Browser.pdf
CONVERTED PDF FILE(S) **(2 pages)**	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0018.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0019.JPG
ORIGINAL PDF FILE	evi_1-207250239110-120501607_._The_Social_Web_Browser.pdf
CONVERTED PDF FILE(S) **(2 pages)**	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0020.JPG
	\\TICRS\EXPORT16\IMAGEOUT16\855\611\85561168\xml6\RFR0021.JPG
DESCRIPTION OF EVIDENCE FILE	pdf's of the term Web

SIGNATURE SECTION

RESPONSE SIGNATURE	/Tye Biasco/
SIGNATORY'S NAME	Tye Biasco
SIGNATORY'S POSITION	Attorney of record, Minnesota bar member
SIGNATORY'S PHONE NUMBER	6123493010
DATE SIGNED	08/26/2013
AUTHORIZED SIGNATORY	YES
CONCURRENT APPEAL NOTICE FILED	YES

FILING INFORMATION SECTION

SUBMIT DATE	Mon Aug 26 14:18:18 EDT 2013
TEAS STAMP	USPTO/RFR-207.250.239.110 -20130826141818264957-855 61168-50046b7bb71ccc3655e 16694f412d767fd964f7c8d28 f13d47fd69458bc7d89-N/A-N /A-20130826120501607891

Request for Reconsideration after Final Action

To the Commissioner for Trademarks:

Application serial no. **85561168** has been amended as follows:

ARGUMENT(S)
In response to the substantive refusal(s), please note the following:

Applicant appreciates the acknowledgment by the Examining Attorney that the amendment to the services listed for International Class 042 and substitute specimen submitted for International Class 042 have satisfied requirement of the earlier office action.

The office action maintains that there is a likelihood of confusion between Applicant's mark and the cited mark because the term WEB INTELLIGENCE is the dominant feature of the two marks and the goods sold under the marks are related. Applicant respectfully requests reconsideration.

The office action indicates that there is potential confusion between the cited mark and Applicant's mark because the goods sold under the two marks are related. The Examining Attorney indicated that the goods associated with the cited mark overlap those of Applicant's goods because the Registrant's goods are not limited to the "surface web." However, the Registrant's goods are limited to "information stored in relational databases," which is the same as the "surface web."

The Examining Attorney also maintains that Applicant's mark of DEEP WEB INTELLIGENCE is merely Registrant's WEB INTELLIGENCE mark with the simple addition of a descriptive term. However, as previously presented by Applicant and further clarified below, the customers of both the Registrant and Applicant understand that the term DEEP WEB has a completely different connotation than WEB with a simple modifier.

A. Applicant's Goods and Services Are Distinct From the Goods of the Cited Mark.

While Applicant's goods and Registrant's goods may be complementary or related, the Office Action misses a significant point. Applicant's goods and services are expressly limited to the " Deep Web." Registrant's goods are expressly limited to " relational databases." These areas of the internet are exclusive to one another. Applicant's goods and services are used for searching data not found in relational databases.

Attached is a whitepaper written by the CEO of Applicant published in March 2013 that explains the significant difference between searching relational databases versus searching the Deep Web. Contrary to the assertion in the office action that the Registrant's mark is not limited to the "surface web," the goods are expressly limited "relational databases," i.e., structured content or, as previously noted, "the surface web."

The goods and services of the parties need not be identical or directly competitive to find a likelihood of confusion. However, they need be related in some manner, or the conditions surrounding their marketing be such, that they could be encountered by the same purchasers under circumstances that could give rise to the mistaken belief that the goods and services come from a common source. *In re Martin's Famous Pastry Shoppe, Inc.*, 748 F.2d 1565, 223 U.S.P.Q. 1289 (Fed. Cir. 1984); *In re Corning Glass Works*, 229 U.S.P.Q. 65 (T.T.A.B. 1985); *In re Rexel Inc.*, 223 U.S.P.Q. 830 (T.T.A.B. 1984); *Guardian Products*

Co., Inc. v. Scott Paper Co., 200 U.S.P.Q. 738 (T.T.A.B. 1978); *In re International Telephone & Telegraph Corp.*, 197 U.S.P.Q. 910 (T.T.A.B. 1978). As outlined in the attached whitepaper, Applicant's goods and services are utilized by military, intelligence, and law enforcement agencies for mining information beyond that available via search engines used with Registrant's goods. Therefore, it is submitted that confusion is unlikely.

B. Cited Marks are Not Distinctive.

Another factor that supports the fact that there is no likelihood of confusion is the fact that there are several marks using the term WEB that have been deemed allowable over each other. Attached is the first page of the results for current registrations using the term WEB in International Class 009 showing there are 724 such registrations. As stated by J. Thomas McCarthy, the noted trademark authority in his well-known treatise Trademarks and Unfair Competition, Fourth Edition, Volume 2, § 11.85:

> A mark that is hemmed in on all sides by similar marks on similar goods cannot be very "distinctive." It is merely one of a crowd of marks. In such a crowd, customers will not be likely to be confused between any two of the crowd. In such a crowd, customers will not likely be confused between any two of the crowd and may have learned to carefully pick out one from the other.

Not only are there numerous registrations in International Class 009 for WEB, but there are examples of what the Examining Attorney contends is a "modifier" used with another mark. Attached are summaries of U.S. Registration No. 3642023 for THE SOCIAL WEB BROWSER and U.S. Registration No. 4303507 for WEB BROWSER, both in International Class 009.

The basic factor in the analysis is whether the relevant purchasing public is likely to be confused by the use of the Applicant's mark. Whereas, herein, there are a number of similar marks for related goods, that the public is not likely to associate the Applicant's product bearing the mark WEB INTELLIGENCE with the Registrant's mark or the other WEB above.

C. Applicant's and Registrant's Respective Customers are Sophisticated and Discerning .

In further support of registration, Applicant notes that the sophistication of the buyer is an important factor in the likelihood of confusion analysis. In *Mead Data Central. Inc. v. Toyota Motor Sales USA. Inc.,* 9 U.S.P.Q. 2d 1442, 1449 (S.D. N.Y.), the Court elaborated on the importance of this factor to the likelihood of confusion analysis, saying that "typically, sophistication of the buyer is a factor that will weigh against finding a likelihood of confusion." As noted above, Applicant's customers are sophisticated, informed purchasers of the military, intelligence, and law enforcement community seeking information that is not found on the surface web. Registrant's customers seek business data in relational databases (see attached brochure of Registrant's product). It is likely that consumers of companies that use relational database searching would not be involved in unstructured data searching or, if so, would understand the differences in such services.

Given the parties' respective clientele, it is clear that confusion is not likely because Applicant's customers are sophisticated enough to distinguish between Applicant's mark and Registrant's mark. Further, Applicant's mark and the cited mark are not likely to be encountered by the same customers because of the Applicant's specialty field of trade. Considering these facts, Applicant submits that confusion is not likely.

The basic factor in the analysis is whether the relevant purchasing public is likely to be confused by the

use of Applicant's mark. Due to the sophistication of Applicant's and Registrant's customers and the exclusivity of the goods for which the marks apply, there is no likelihood of confusion.

Applicant accordingly submits that the objections should be withdrawn and its mark approved for publication. The Examiner is invited to telephone the undersigned if the Examiner believes it would be useful to advance prosecution.

EVIDENCE
Evidence in the nature of pdf's of the term Web has been attached.
Original PDF file:
evi_207250239110-120501607_._Deep_Web_Whitepaper.pdf
Converted PDF file(s) (11 pages)
Evidence-1
Evidence-2
Evidence-3
Evidence-4
Evidence-5
Evidence-6
Evidence-7
Evidence-8
Evidence-9
Evidence-10
Evidence-11
Original PDF file:
evi_1-207250239110-120501607_._List_of_WEB_Registrations.pdf
Converted PDF file(s) (5 pages)
Evidence-1
Evidence-2
Evidence-3
Evidence-4
Evidence-5
Original PDF file:
evi_2-207250239110-120501607_._Web_Browser.pdf
Converted PDF file(s) (2 pages)
Evidence-1
Evidence-2
Original PDF file:
evi_1-207250239110-120501607_._The_Social_Web_Browser.pdf
Converted PDF file(s) (2 pages)
Evidence-1
Evidence-2

SIGNATURE(S)
Request for Reconsideration Signature
Signature: /Tye Biasco/   Date: 08/26/2013
Signatory's Name: Tye Biasco
Signatory's Position: Attorney of record, Minnesota bar member

Signatory's Phone Number: 6123493010

The signatory has confirmed that he/she is an attorney who is a member in good standing of the bar of the highest court of a U.S. state, which includes the District of Columbia, Puerto Rico, and other federal territories and possessions; and he/she is currently the applicant's attorney or an associate thereof; and to the best of his/her knowledge, if prior to his/her appointment another U.S. attorney or a Canadian attorney/agent not currently associated with his/her company/firm previously represented the applicant in this matter: (1) the applicant has filed or is concurrently filing a signed revocation of or substitute power of attorney with the USPTO; (2) the USPTO has granted the request of the prior representative to withdraw; (3) the applicant has filed a power of attorney appointing him/her in this matter; or (4) the applicant's appointed U.S. attorney or Canadian attorney/agent has filed a power of attorney appointing him/her as an associate attorney in this matter.

The applicant is filing a Notice of Appeal in conjunction with this Request for Reconsideration.

Serial Number: 85561168
Internet Transmission Date: Mon Aug 26 14:18:18 EDT 2013
TEAS Stamp: USPTO/RFR-207.250.239.110-20130826141818
264957-85561168-50046b7bb71ccc3655e16694
f412d767fd964f7c8d28f13d47fd69458bc7d89-
N/A-N/A-20130826120501607891

UNDERSTANDING THE DEEP WEB IN 10 MINUTES

Learn where it's at, how you can search it, what you'll find there, and why Google can't find everything

by Steve Pederson, CEO
spederson@brightplanet.com
March 2013

BrightPlanet®
deep web intelligence

Understanding the Deep Web in 10 Minutes
Learn where it's at, how you can search it, what you'll find there, and why Google can't find everything

by Steve Pederson

I. Introduction

Don't worry if you don't understand what the term "Deep Web" means. "Deep Web" is a vague description of the internet not necessarily accessible to search engines. The Deep Web is often misinterpreted as the "Dark Web". While browsing the internet, the Deep Web is usually right in front of you, you may just not know it yet. Whether you are searching for unstructured Big Data or trying to answer narrowly targeted questions , it can typically be found somewhere within the millions of Deep Web sources.

Both public and private sector organizations are intrigued by the vast potential of harvesting unstructured content at scale from the internet, tagging entities in the metadata, and curating that semi-structured content into actionable intelligence. There are many questions frequently asked about the process and possibilities for Deep Web harvesting, analytics, and data output. BrightPlanet hopes to answer those common questions in this whitepaper.

In this whitepaper you will discover:

- ✔ Where the Deep Web is, and how it compares to the Surface Web and Dark Web
- ✔ Why you should care about the Deep Web
- ✔ The difference between a search engine and a Deep Web harvest engine
- ✔ How data is harvested from the Deep Web
- ✔ Deep Web harvest use cases

II. What is the Deep Web?

DEEP WEB VS. SURFACE WEB

The Deep Web is a part of the internet not accessible to link-crawling search engines like Google. The only way a user can access this portion of the internet is by typing a directed query into a web search form, thereby retrieving content within a database that is not linked. In layman's terms, the only way to access the Deep Web is by conducting a search that is within a particular website.

The Surface Web is the internet that can be found via link-crawling techniques; link-crawling means linked data can be found via a hyperlink from the homepage of a domain. Google can find this Surface Web data.

Surface Web search engines (Google/Bing/Yahoo!) can lead you to websites that have unstructured Deep Web content. Think of searching for government grants; most researchers start by searching "government grants" in Google, and find few specific listings for government grant sites that contain databases. Google will direct researchers to the website www.grants.gov, but not to specific grants within the website's database.

Researchers can search thousands of grants at www.grants.gov by searching the database via the website search box. In this example, a Surface Web search engine (Google) led users to a Deep Web website (www. grants.gov) where a directed query to the search box brings back Deep Web content not found via Google search.

DARK WEB AND DEEP WEB - NOT THE SAME THING!

The Dark Web refers to any web page that has been concealed to hide in plain sight or reside within a separate, but public layer of the standard internet.

The internet is built around web pages that reference other web pages; if you have a destination web page which has no inbound links you have concealed that page and it cannot be found by users or search engines. One example of this would be a blog posting that has not been published yet. The blog post may exist on the public internet, but unless you know the exact URL, it will never be found.

Other examples of Dark Web content and techniques include:

- Search boxes that will reveal a web page or answer if a special keyword is searched. Try this by searching "distance from Sioux Falls to New York" on Google.
- Sub-domain names that are never linked to; for example, "internal.brightplanet.com"
- Relying on special HTTP headers to show a different version of a web page
- Images that are published but never actually referenced, for example "/image/logo_back.gif"

Virtual private networks are another aspect of the Dark Web that exists within the public internet, which often requires additional software to access. TOR (The Onion Router) is a great example. Hidden within the public web is an entire network of different content which can only be accessed by using the TOR network.

While personal freedom and privacy are admirable goals of the TOR network, the ability to traverse the internet with complete anonymity nurtures a platform ripe for what is considered illegal activity in some countries, including:

- Controlled substance marketplaces
- Armories selling all kinds of weapons
- Child pornography
- Unauthorized leaks of sensitive information
- Money laundering
- Copyright infringement
- Credit Card fraud and identity theft

Users must use an anonymizer to access TOR Network/Dark Web websites. The Silk Road, an online marketplace/infamous drug bazaar on the Dark Web, is inaccessible using a normal search engine or web browser.

WHY SHOULD YOU CARE ABOUT THE DEEP WEB?

For 2013, it is important to tap into the rich resources existing in the Deep Web. The last time an extensive study was completed estimating the size of the Deep Web was in 2001 — a time when the internet consisted of only approximately three million different domains. The 2001 study revealed that at that time the Deep Web was approximately 400-500 times the size of the Surface Web.

Today's internet is significantly bigger with an estimated 555 million domains, each containing thousands

or millions of unique web pages. As the web continues to grow, so too will the Deep Web and the value attained from Deep Web content.

III. Search Engines vs. Deep Web Harvest Engines

Harvesting is the term BrightPlanet uses when it talks about accessing the Deep Web. It is important to distinguish between traditional searches and Deep Web harvesting. Unlike traditional search technologies, like Google, that index links and allow you to view the results, BrightPlanet takes it a step further and harvests all of the results. The harvest process involves BrightPlanet extracting all of the text based content from each of the results pages and then preparing the content for some type of analysis depending on the needs of end-users.

To understand the major differences between a harvest engine and a search engine, it's important to understand the problem that search engines are meant to solve.

YESTERDAY'S SEARCH ENGINES

The problem search engines tried to tackle dates back to the early 1990s as the internet increased in popularity. Mostly static web pages were being added to the internet, but users needed a way to easily find web pages that contained information.

Search engines like Google, AltaVista, Yahoo!, and Lycos created technologies that crawled through websites and indexed them as a way for users to identify pages of interest. Search engines tried to find the most relevant page containing the answer to what users were looking for.

Questions that were originally asked to search engines in the late 90's were very basic. Students researching class reports replaced encyclopedias with the internet, researchers created basic web pages to share their discoveries, and social sharing consisted of updating your GeoCities page. The 90's internet was non-commercial and viewed with a research purpose.

TODAY'S SEARCH ENGINES

Today's internet is significantly different; millions of web pages are published for all sorts of reasons beyond traditional research.

Search engine companies developed systems able to quickly index millions of web pages in a short time period, therefore allowing users to accurately search the assimilated index. Search engines don't find or store all the content on a web page; they simply lead you to the content's location. This lack of data retention allows search engines to get away with storing minimal information about each individual web page.

Typically, search engines store the most frequently mentioned words, locations of those words, and any metadata (title of the web page, URL of the web page, keywords, etc) when indexing web pages. **The amount of data stored from each page is a crucial difference between search engines and harvesters.**

SEARCH ENGINES AND THE SURFACE WEB

Search engines like Google are really good at finding Surface Web websites; providing answers to basic questions quickly. However, companies and organizations have significantly harder questions than "How late is Burger King open?" Complex questions like those listed below require more than a search engine; they require a Deep Web Harvester®:

- Who is selling my products fraudulently online?
- How many people have won grants on Fetal Alcohol Spectrum Disorders?
- What are clinical trial patients saying about my experimental drug?
- What new information has been published on my competitor's website today?
- Has anything changed in this insurance coverage plan that would affect a pharmaceutical company's stock price?
- What new breast cancer research has been published in the last month? What are people saying about it?

DEEP WEB HARVEST ENGINE

Unlike a search engine, BrightPlanet's Deep Web Harvester extracts every single word every time it accesses a web page. Additionally, the Deep Web Harvester stores every single page harvested as a separate version in its database.

For example, BrightPlanet has a list of 100 websites actively harvesting for a customer every four hours. Therefore, the Deep Web Harvester collects a version of every single web page found within the 100 domains every four hours.

To put that into perspective, let's envision that each of those domains is a relatively small website (100 pages). In this scenario, every four hours we harvest content from 10,000 web pages (100 web pages multiplied by 100 domains). In one week, this harvesting process stores 420,000 web pages. BrightPlanet harvested 53 million web pages over a 30-day period for one customer.

A. DEEP WEB HARVEST ADVANTAGES

The concept of a harvest engine has a number of different advantages. The two largest advantages being:

- Analytic capabilities
- Versioning of web pages

Because BrightPlanet harvests the actual raw text from web pages, as opposed to storing metadata and only top keywords, BrightPlanet can integrate its harvested data directly into nearly any analytic technology using our OpenPlanet® Enterprise Platform [see page 7 for more on OpenPlanet].

Combining BrightPlanet's scalable harvesting capabilities with custom analytic technology helps customers visualize, analyze, and ultimately create intelligence from large data sets.

IV. Where do Deep Web websites come from?

SOURCE REPOSITORY: A LIBRARY OF 85,000 (AND GROWING) DEEP WEB SOURCES

The Deep Web is at least 400-500 times the size of the Surface Web. It is continuously growing, and that means new Deep Web sources are also growing. BrightPlanet harnesses Deep Web sources by sorting and indexing them in its Source Repository.

The Source Repository is a library of Deep Web sources/websites that BrightPlanet has collected over 10 years of web harvesting on behalf of clients.

New sources are added and updated daily. There are currently over 85,000 Deep Web sources, grouped

by source type, in BrightPlanet's Source Repository. Examples of source type groups include Law, Healthcare, Pharmaceuticals, Social Media, Major Media, Newspapers, Finance & Economics, and Politics to name just a handful of the over 60 groups.

HOW YOU CAN LEVERAGE THE SOURCE REPOSITORY

End-users do not need to worry about communication with sources; those processes are all done automatically by BrightPlanet. You just need to identify the information you are trying to find and from what sources, and BrightPlanet can harvest it on your behalf.

BrightPlanet commonly works with its end-users to harvest content from custom Deep Web sources. End-users define hundreds or thousands of Deep Web sources for BrightPlanet to query with many keywords at once. Once new sources are entered into the Source Repository, they will be indexed and saved for future harvests.

Here are just a few examples of how the Source Repository can be leveraged:

Newspapers

The Newspapers group in the Source Repository includes every newspaper in the U.S. In a matter of seconds, BrightPlanet could harvest topic specific content from every newspaper in the U.S. Instead of searching newspaper website after newspaper website, the information could be harvested instantly. Additionally, the papers are sorted by state so you could limit the search to certain states if it better fits your needs.

Law

There are several categories within this group. One of those categories is Courts. This group includes sources that would allow you to search Court rulings at all levels of the judicial branch; state, local, and federal, instantly.

Finance & Markets

Buy the rumor; sell the news. Users can find both rumors and news faster than the competition by harvesting from the News, Finance Blog/Website, Finance Message Board, and industry-specific blogs and message board source groups.

Health & Pharmaceutical

There are dozens of possibilities for leveraging the Source Repository for the health and pharmaceutical sectors: fraud, diversion, health websites, disease-specific websites, and message boards to name a few. For example, if you wanted to search for any mentions of a new multiple sclerosis drug, selecting the M.S. Message Boards, M.S. Blogs/Websites, Health Blogs/Websites, and Health Message Boards source groups yields access to 75 reliable Deep Web sources for you to instantly search.

VIEWING THE CONTENT IN DEEP WEB INTEL SILOS

Another solution BrightPlanet offers, to help sort and view harvested Deep Web data, is Deep Web Intel Silos. For the purposes of this paper, we'll talk about how healthcare research has leveraged Deep Web Intel Silos.

There are millions of documents available on the Deep Web for healthcare research that current methods of online research have no way of finding or collecting. Deep Web Intel Silos can create collections of nearly any open-source content. For healthcare research, BrightPlanet creates disease and healthcare

topic-specific research silos to which researchers subscribe.

Unlike a traditional static database like PubMed or LexisNexis, where the dataset is predefined by the organization offering access, topic-specific research silos start with a base set of data and add additional sources requested by subscribers. This allows for collaboration between research institutions.
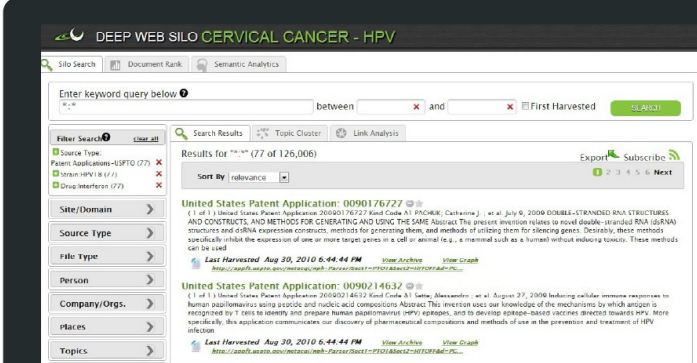
As more and more researchers request sources to be added to research silos, and as BrightPlanet continuously monitors these sources in key topic areas, research silos develop into some of the most comprehensive topic-specific research databases worldwide.

Since the subject matter experts, healthcare researchers in this case, identify the sources and source types they want to draw from, and dictate how they want harvested documents tagged and sorted. Tagging documents becomes crucial when creating intelligence from large datasets; the big challenge everyone has with Big Data. The final Deep Web Intel Silo dataset contains only relevant, searchable data with customizable drill-down search facets.

REFINING A LARGE SET OF RELEVANT DATA INTO ACTIONABLE INTELLIGENCE

Let's say a research silo contains 126,000 harvested documents related to the broad topic of cervical cancer. If the researcher is only interested in patent applications mentioning the drug Interferon with the HPV18 strain, the user can create an advanced search focused only on patent applications.

By narrowing the search to only patent applications, the huge dataset is reduced to 77 relevant patent applications mentioning HPV18 and the drug Interferon. Any additional search queries the user performs will comb through only those 77 super-relevant documents.

**Source Type: Patent Applications
Strain: HPV18
Drug: Interferon**

THE VALUE OF SILO SERVICES

Deep Web Intel Silos are individual repositories for topic-specific content, and are updated with new and relevant information from the harvester in real-time. Each silo is filled with high-quality Deep Web resources – databases, RSS feeds, and more – that lie beyond the reach of traditional search engines. They also include standard analytical tools like raw data views, topic clustering, and link analysis. Additional custom analytical modules can easily be added to meet your reporting needs.

The true value of silo services lies with BrightPlanet's Deep Web Researchers. These are highly skilled content managers who take the complexity out of Deep Web research. Think of them as your personal guides to the Deep Web, discovering and harvesting the resources that fill your silo with relevant, timely content. Our researchers work hard to deliver the best results available, leaving you time to do what you do best: analyze, interpret, and create actionable intelligence.

LARGE SCALE – OPENPLANET

Many customers only require access to harvested content to make searching capabilities simpler for them. For these specific customers, access to a Deep Web Intel Silos fulfills their needs. Customers wanting to make additional conclusions from harvested data can easily integrate the data into any number of analytic capabilities.

Through its tenure with the U.S. intelligence community, BrightPlanet has learned that a single end-to-end harvest platform takes anywhere from six months to three years to set up, depending on the scale and number of components. While this is a good business opportunity for system integrators who can bill hourly, it is not a desirable solution for commercial deployments demanding a higher level of integration without custom development. BrightPlanet saw this need for open integration early on and spent two years developing an open platform called OpenPlanet to overcome these limitations.

The OpenPlanet platform is based on a simple workflow that completely separates the harvesting and analytic components of data collection and analysis. This concept allows BrightPlanet to easily swap in and out different analytic technologies with no knowledge of where the data comes from. Allowing customers to integrate multiple datasets, not just harvested web data, with multiple analytic technologies in one workflow without significant development.
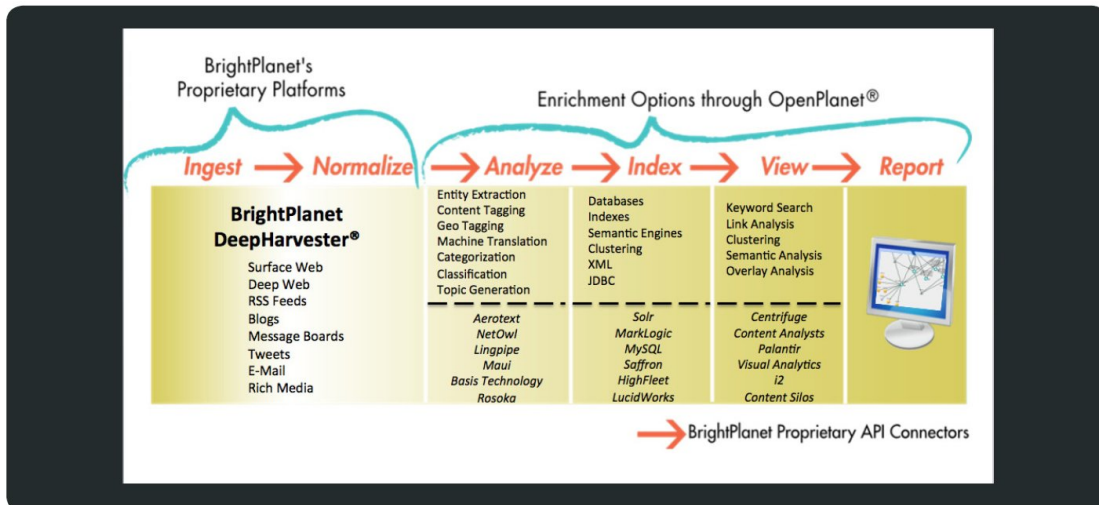

V. Deep Web Harvest End-User Examples

A DEEP WEB HARVEST VS. SEARCH ENGINE USE CASE

The following example shows the kind of Deep Web content search engines may be missing.

The *Argus Leader*, the local newspaper of Sioux Falls, South Dakota, did an article about BrightPlanet's CEO, Steve Pederson (an avid bugler) titled "Living Legacy in 24 Notes." The article at one point in time had been on the homepage of the Argus Leader, a location that is reachable by a Surface Web search engine like Google. A few days after the article was featured on the homepage, the article was pushed into archive format, and thus only reachable via a query through the search box located on Argus Leader's site; it left the Surface Web and entered the Deep Web.

These two images demonstrate the differences between the Deep Web and the Surface Web. The image on the left is a search of what Google has indexed. The query (BrightPlanet AND "Steve Pederson" site:argusleader.com) tells Google that the only results we want are from the Argus Leader domain. The search returns zero web pages that have been indexed by Google containing both BrightPlanet AND "Steve Pederson".
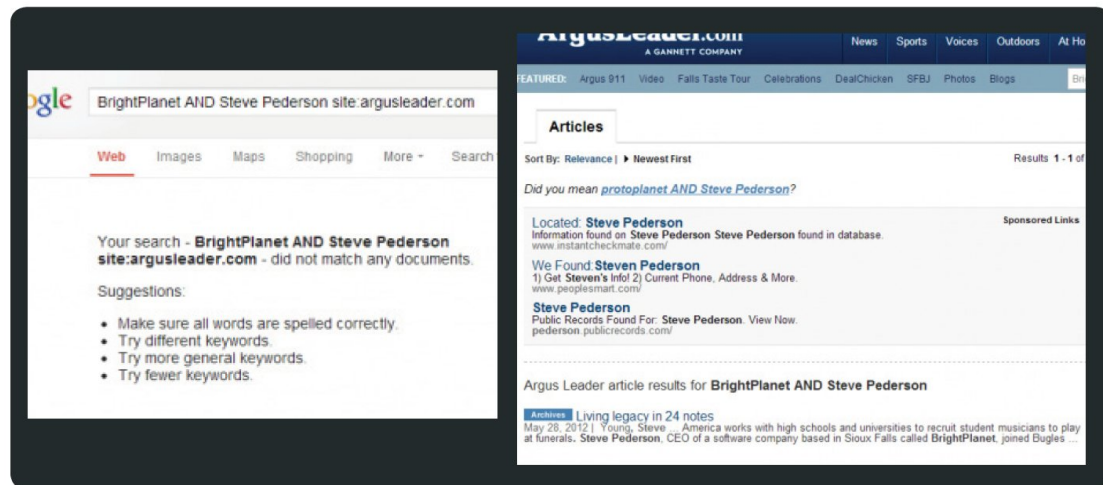
The image on the right proves that results containing both terms do exist. This search is performed using the search box provided by the *Argus Leader* website. The reason why this search returns results is because the search box points to the newspaper's database, a Deep Web source. Archived content can only be accessed via the Argus Leader's search, making that content exclusive to the Deep Web. Google does not direct queries into any site searches, as it only finds documents via link following. The "Living Legacy in 24 Notes" news article has fallen into the Deep Web.

When BrightPlanet collects Deep Web content, it is exactly this type of search placed directly into the search forms that BrightPlanet can execute at a very large scale; issuing thousands of search queries into thousands of Deep Web sites and pulling all the content back for analysis. Imagine being able to query every single online newspaper web search form within the United States simultaneously.
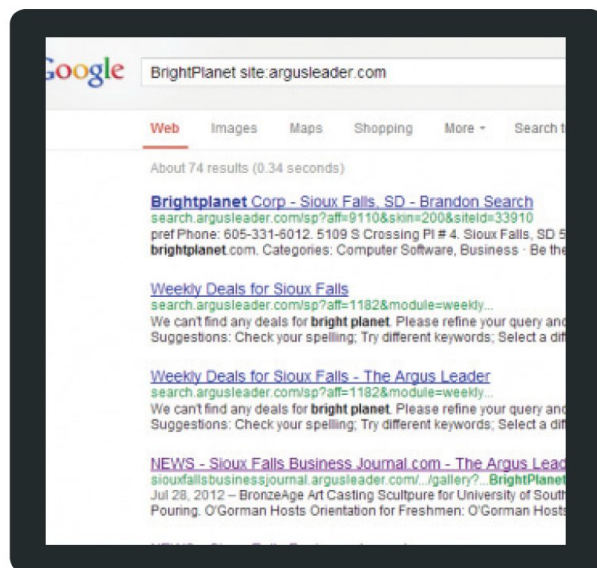
The other major advantage of using a Deep Web harvest over a search engine is efficiency. Doing a search for the query BrightPlanet on the Argus Leader web page will return the same one article. Doing a search for BrightPlanet within the Argus Leader domain on Google will return 74 results (see image on left).
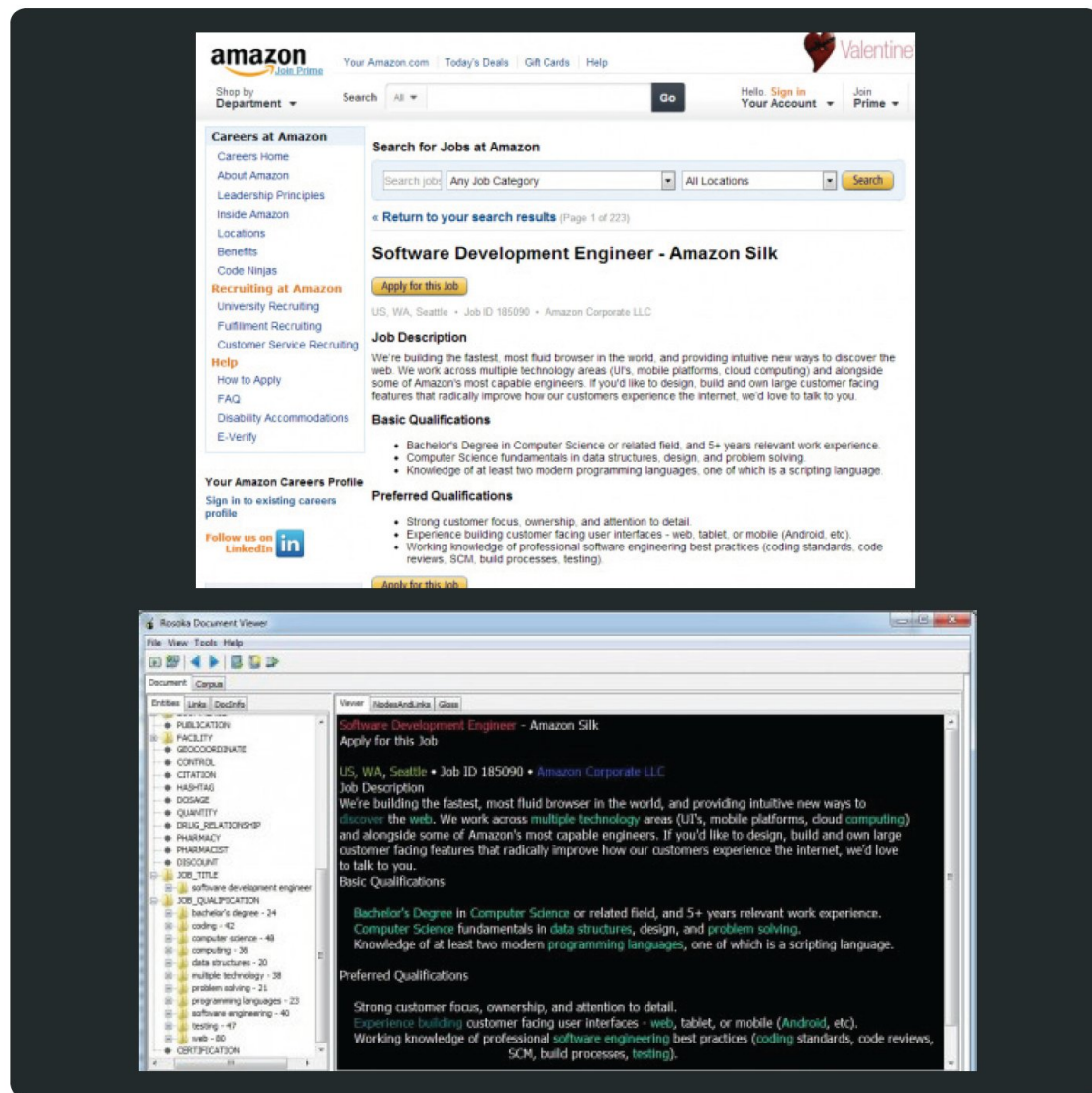
The extra 73 results return links that no longer contain BrightPlanet on the actual page, as Google is still searching an old version of the page. When Google crawls through a site, it filters through millions of links, often picking up irrelevant content. When BrightPlanet performs a Deep Web search on a site, it only harvests the relevant content related to your queries.

AN OPENPLANET USE CASE

An interesting example of following data from the harvest stage through the OpenPlanet Platform is a recent project BrightPlanet completed for a management consulting firm; BrightPlanet delivered bi-weekly exports of all job postings from every Fortune 200 company.

First, BrightPlanet harvested all job postings from Fortune 200 companies. The raw text of each job posting wasn't enough to give insight into the hiring actions of the companies, so BrightPlanet worked with the end-user to enrich the content with custom tags. BrightPlanet wrote custom rules and dynamically tagged and extracted the locations of the job postings, job titles, job qualifications, and required certifications.

The deliverable for the end-user was a CSV file consisting of the company, job title, location, important qualifications, URL, and the raw text of each posting. The end-user uploaded the data into their own database for analysis, and the management consulting firm was able to add value to its product offerings.

The images on page 9 show a web page (top) and what it looks like once it is normalized and the entities are extracted (bottom). The image on the bottom is displaying the web page in Rosoka's Document viewer. The highlighted text in the second image displays entities that have been extracted from the text of the job posting.

Even though the final deliverable for the end-user was not an analytic interface or report, it's easy to see the insight you could quickly draw from the job posting output. For example, users could quickly identify which companies had postings for computer programmers that require Java programming skills. Many valuable insights can be drawn from the data set because of the extracted and enriched data BrightPlanet provided.

VI. Conquer Big Data by Pairing Internal Data with Unstructured Deep Web Data

Big Data doesn't just come from within company walls. Structured enterprise data is only one part of the hybrid data spectrum. Unstructured data found on the internet is the other part, and there are trillions of unstructured web documents ready to be harvested.

Search engines are a good starting point, but they only skim the surface of available content. The more-efficient source qualification and filtering capabilities of Deep Web harvests lead to less time searching, leaving more time for creating actionable intelligence.

Now that you know what the Deep Web is and the many ways to get data from it, what actionable intelligence can BrightPlanet help you harvest from the Deep Web? The Deep Web grows exponentially every year; start tapping into it for your business or institution.

Trademarks > Trademark Electronic Search System (TESS)

TESS was last updated on Tue Aug 6 03:20:30 EDT 2013

| TESS HOME | NEW USER | STRUCTURED | FREE FORM | BROWSE DICT | SEARCH OG | PREV LIST | NEXT LIST | IMAGE LIST | BOTTOM | HELP |

Logout | *Please logout when you are done to release system resources allocated for you.*

Start | List At: [   ] OR Jump to record: [   ] **724 Records(s) found (This page: 1 ~ 100)**

Refine Search [(*web*)[MI] and (009)[IC] and (LIVE)[LD] and `] Submit

Current Search: (*web*)[MI] and (009)[IC] and (LIVE)[LD] and `RN > "0" docs: 724 occ: 2900

	Serial Number	Reg. Number	Word Mark	Check Status	Live/Dead
1	85140093	4376825	MAKING THE WEB MOBILE	TSDR	LIVE
2	85580269	4373330	WILEY WEBB	TSDR	LIVE
3	85626383	4304296	WEBTHORITY	TSDR	LIVE
4	85681467	4272400	GO DADDY WEBSITE BUILDER	TSDR	LIVE
5	85449293	4359844	THE WEBCOMMUNITY	TSDR	LIVE
6	85112987	4358351	WEBBO	TSDR	LIVE
7	85737770	4309205	WEBGREASE	TSDR	LIVE
8	85635976	4344386	WEBS	TSDR	LIVE
9	85404883	4343211	INFINITE VISIONS WEB EDITION	TSDR	LIVE
10	85402668	4340718	SMART WEB	TSDR	LIVE
11	85725055	4338067	WEBGILITY	TSDR	LIVE
12	85483568	4326676	NOTES2WEB	TSDR	LIVE
13	85298526	4328791	THE VOICE OF THE WEB	TSDR	LIVE
14	85104150	4313438	WEBMEUP	TSDR	LIVE

15	85570686	4310486	WEBSTER	TSDR	LIVE
16	85519916	4310306	MYWEB	TSDR	LIVE
17	85272290	4303507	WEB BROWSER	TSDR	LIVE
18	85501177	4300103	WEBSITEFORGE	TSDR	LIVE
19	85634980	4288381	RC-WEBVIEW	TSDR	LIVE
20	85529962	4299215	SLIM-WEBSERVICES	TSDR	LIVE
21	85482423	4299095	WEBSIM	TSDR	LIVE
22	85472821	4299040	WEBREP	TSDR	LIVE
23	85041789	4296939	WEBKIT	TSDR	LIVE
24	85676855	4295728	WEBFIRE	TSDR	LIVE
25	85676797	4295727	WEBFIRE	TSDR	LIVE
26	85395835	4296226	WEBROOT SECUREANYWHERE	TSDR	LIVE
27	85395815	4296224	WEBROOT SECUREANYWHERE	TSDR	LIVE
28	85672229	4292312	MOOVWEB	TSDR	LIVE
29	85665650	4289000	SIGNWEB	TSDR	LIVE
30	85665609	4288999	SCREENWEB	TSDR	LIVE
31	85644732	4287063	WEB SUPPLIER LABELING	TSDR	LIVE
32	85513835	4232845	CFOWEB.COM	TSDR	LIVE
33	85224411	4265075	WEBYDO	TSDR	LIVE
34	85413056	4258703	DYNAMICWEB	TSDR	LIVE
35	85342779	4258508	ATHEROWEB	TSDR	LIVE
36	85123595	4047862	WEBTRENDS SEGMENTS	TSDR	LIVE
37	85551022	4253019	WEBCAST METRICS	TSDR	LIVE
38	85492847	4249003	UC WEB	TSDR	LIVE
39	85977437	4247531	THE WEB IS WHAT YOU MAKE OF IT	TSDR	LIVE
40	85015219	4008265	WHAT'S WHEN ON THE WEB	TSDR	LIVE
41	85234915	4230622	KING OF THE WEB	TSDR	LIVE
42	85145515	3988870	AIMS WEBLINK	TSDR	LIVE
43	85547588	4221471	THE WEB'S ORIGINAL VOICE OF WHOLESALE	TSDR	LIVE
44	85542190	4209667	WEBHOTEL	TSDR	LIVE
45	85207536	4196669	WEBSCORE	TSDR	LIVE

46	85445728	4189485	WEB2IMPRESS	TSDR	LIVE
47	85470905	4171405	WEBSCORER	TSDR	LIVE
48	85200149	3993807	WEBIC	TSDR	LIVE
49	85314593	4173927	SWEB	TSDR	LIVE
50	85344021	4170406	WEBMOBI	TSDR	LIVE
51	85459294	4164436	WEBOP	TSDR	LIVE
52	85389732	4163664	WEBSTORM	TSDR	LIVE
53	85467847	4161365	WEBBER	TSDR	LIVE
54	85026324	4068085	WE MAKE WEBSITES EASY	TSDR	LIVE
55	85415642	4141984	STEADY WEB	TSDR	LIVE
56	85256964	4141142	ONCOWEB	TSDR	LIVE
57	85416891	4138505	INDUSOFT WEB STUDIO (IWS)	TSDR	LIVE
58	85141819	4140433	AVWEBAUDIO	TSDR	LIVE
59	85325474	4137155	WEB DESIGN AUTOMATION	TSDR	LIVE
60	85248252	4136694	WEBCX	TSDR	LIVE
61	85060376	4130426	ICEWEB	TSDR	LIVE
62	85284633	4129749	WEBFORMZ FOR IMIS	TSDR	LIVE
63	85382611	4125117	DEALER WEB WATCHER	TSDR	LIVE
64	85231571	4093227	WEBINSPECT	TSDR	LIVE
65	85195037	3993523	WEBTAS	TSDR	LIVE
66	85187232	4086396	DOCULEX WEBSEARCH	TSDR	LIVE
67	85201091	4082388	WEBTASFX	TSDR	LIVE
68	85347492	4074292	GFI WEBMONITOR	TSDR	LIVE
69	85352511	4072248	WEBCOLLECT	TSDR	LIVE
70	85106555	4069443	WEB RACING	TSDR	LIVE
71	85196389	4058851	DEAD ON THE WEB	TSDR	LIVE
72	85034312	4054227	INTELLIGENT WEB CLIENT	TSDR	LIVE
73	85285113	4048799	SPATIALWEB	TSDR	LIVE
74	85975315	4045576	WEBTAS	TSDR	LIVE
75	85251061	4034613	WEBTOP BPM	TSDR	LIVE
76	85065942	4036483	CDIP.WEB	TSDR	LIVE

77	85088643	4024924	XTEND2WEB	TSDR	LIVE
78	85030744	4024784	WEBS2AT	TSDR	LIVE
79	85188569	4021918	FAKE WEBCAM	TSDR	LIVE
80	85160069	4018475	VASONT WEBLNX	TSDR	LIVE
81	85233486	4016040	WEBCATI	TSDR	LIVE
82	85150943	4013931	WEBFOCUS ACTIVE TECHNOLOGIES	TSDR	LIVE
83	85160092	4005299	WEB MEETINGS THAT WOW	TSDR	LIVE
84	85066396	4001099	WEB WATCHER	TSDR	LIVE
85	85150518	3996629	WEBAV	TSDR	LIVE
86	85204770	3994018	MONAHRQ INPUT YOUR DATA.OUTPUT YOUR WEBSITE.	TSDR	LIVE
87	85067488	3992346	WEBASSIGN	TSDR	LIVE
88	85066042	3987074	WEB TRNS PORT	TSDR	LIVE
89	85177951	3966447	MERGENT WEBREPORTS	TSDR	LIVE
90	85024707	3967722	WEBMATRIX	TSDR	LIVE
91	85021974	3964826	WEBCASS	TSDR	LIVE
92	85075197	3950299	QUICKCHECK WEBLINE	TSDR	LIVE
93	85068872	3911204	WEBTRAC	TSDR	LIVE
94	85052268	3907492	EDUWEB	TSDR	LIVE
95	85018482	3889635	WEB COMFORT	TSDR	LIVE
96	85015148	3921207	WEBRADAR	TSDR	LIVE
97	85006053	3876662	WE UP EVERY WEBSITE VISITOR YOU SELL MORE CARS! CLIENT CONNEXION WWW.CLIENTCONNEXION.COM	TSDR	LIVE
98	79023414	3256664	VISIWEB	TSDR	LIVE
99	79023246	3256657	VISIWEB	TSDR	LIVE
100	79100463	4277679	WEBRECRUIT	TSDR	LIVE

Trademarks > Trademark Electronic Search System (TESS)

TESS was last updated on Tue Aug 6 03:20:30 EDT 2013

TESS HOME | NEW USER | STRUCTURED | FREE FORM | BROWSE DICT | SEARCH OG | BOTTOM | HELP | PREV LIST | CURR LIST | NEXT LIST | FIRST DOC | PREV DOC | NEXT DOC | LAST DOC

Logout   Please logout when you are done to release system resources allocated for you.

Start   List At: [   ]   OR   Jump   to record: [   ]   **Record 17 out of 724**

TSDR   ASSIGN Status   TTAB Status   *(Use the "Back" button of the Internet Browser to return to TESS)*

Word Mark	WEB BROWSER
Goods and Services	IC 009. US 021 023 026 036 038. G & S: Television receivers
Mark Drawing Code	(3) DESIGN PLUS WORDS, LETTERS, AND/OR NUMBERS
Design Search Code	01.07.01 - Globes with outlines of continents
Serial Number	85272290
Filing Date	March 21, 2011
Current Basis	44E

Original Filing Basis	1B;44D
Published for Opposition	January 1, 2013
Registration Number	**4303507**
Registration Date	March 19, 2013
Owner	(REGISTRANT) LG Electronics Inc. CORPORATION REPUBLIC OF KOREA 20 Yeouido-dong Youngdeungpo-gu Seoul 150-721 REPUBLIC OF KOREA
Attorney of Record	Robert J. Kenney
Priority Date	February 11, 2011
Disclaimer	NO CLAIM IS MADE TO THE EXCLUSIVE RIGHT TO USE "WEB BROWSER" APART FROM THE MARK AS SHOWN
Description of Mark	The color(s) blue, gray and red is/are claimed as a feature of the mark. The mark consists of a globe in blue and gray with "WEB" written in gray to the right of the globe with "BROWSER" in red underneath "WEB".
Type of Mark	TRADEMARK
Register	PRINCIPAL
Live/Dead Indicator	**LIVE**

Trademarks > Trademark Electronic Search System (TESS)

TESS was last updated on Tue Aug 6 03:20:30 EDT 2013

| TESS HOME | NEW USER | STRUCTURED | FREE FORM | BROWSE DICT | SEARCH OG | BOTTOM | HELP | PREV LIST | CURR LIST | NEXT LIST | FIRST DOC | PREV DOC | NEXT DOC |
| LAST DOC |

Logout   Please logout when you are done to release system resources allocated for you.

Start   List At: [   ] OR Jump   to record: [   ]   **Record 424 out of 724**

TSDR   ASSIGN Status   TTAB Status   *(Use the "Back" button of the Internet Browser to return to TESS)*

THE SOCIAL WEB BROWSER

Word Mark	THE SOCIAL WEB BROWSER
Goods and Services	IC 009. US 021 023 026 036 038. G & S: Internet web browser software; software for accessing information and sites on a global computer network; communications software for connecting global computer networks, software for the transmission and access of electronic mail, instant messaging, voice over IP telephony, online blog journals, text, weblinks, images, audio and video content over a global computer network; computer software used to enhance the capabilities and features of other software. FIRST USE: 20061100. FIRST USE IN COMMERCE: 20061100
	IC 038. US 100 101 104. G & S: online journals, namely, blogs in the field of general interest; electronic transmission of data and documents over the Internet or other communications network. FIRST USE: 20061100. FIRST USE IN COMMERCE: 20061100
Standard Characters	

Claimed Mark Drawing Code	(4) STANDARD CHARACTER MARK
Serial Number	77461533
Filing Date	April 30, 2008
Current Basis	1A
Original Filing Basis	1A
Published for Opposition	April 7, 2009
Registration Number	**3642023**
Registration Date	June 23, 2009
Owner	(REGISTRANT) Flock, Inc. CORPORATION DELAWARE 805 Veterans Blvd., Suite 307 Redwood City CALIFORNIA 94063
Attorney of Record	David A.W. Wong
Disclaimer	NO CLAIM IS MADE TO THE EXCLUSIVE RIGHT TO USE "WEB BROWSER" APART FROM THE MARK AS SHOWN
Type of Mark	TRADEMARK. SERVICE MARK
Register	PRINCIPAL
Live/Dead Indicator	**LIVE**

TESS HOME | NEW USER | STRUCTURED | FREE FORM | BROWSE DICT | SEARCH OG | TOP | HELP | PREV LIST | CURR LIST | NEXT LIST | FIRST DOC | PREV DOC | NEXT DOC | LAST DOC

|.HOME | SITE INDEX| SEARCH | eBUSINESS | HELP | PRIVACY POLICY